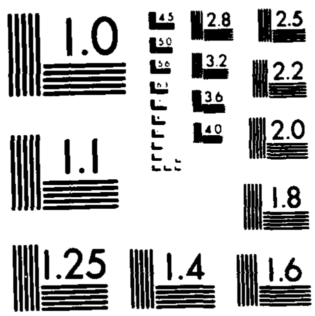
AD-A130 705 WEAK AND STRONG LAW RESULTS FOR A FUNCTION OF THE SPACINGS(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PRECESSES W P MCCORMICK MAY 83 TR-30 1/1

UNCLASSIFIED AFOSR-TR-83-0627 F49620-82-C-0009 F/G 12/1 NL

END
DATE
FILMED
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

(2)

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

WEAK AND STRONG LAW RESULTS FOR A FUNCTION OF THE SPACINGS

William P. McCormick

TECHNICAL REPORT #30

May 1983

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0627	2. GOVT ACCESSION NO. AD-A130 705	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) WEAK AND STRONG LAW RESULTS FOR A FUNCTION OF THE SPACINGS		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William P. McCormick		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0009
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of North Carolina Chapel Hill, NC 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR /NM Bolling AFB Washington, DC 20332		12. REPORT DATE May 1983
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release-- distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Order Statistic, spacings, limiting distribution, strong law.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $\{U_n, n\geq 1\}$ be i.i.d. uniform on $(0,1)$ random variables and define $S_{i,n} = U_{i,n-1} - U_{i-1,n-1}$, $i=1,\ldots,n$ where the $U_{i,n-1}$ are the order statistics from a sample of size $n-1$ and $U_{0,n-1} = 0$ and $U_{n,n-1} = 1$. The $S_{i,n}$ are called the spacings divided by $U_1,\ldots,U_{n-1}$. For a fixed integer ℓ, set $M_{\ell,n} =$

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

UNCLASSIFIED

$\max_{1 \le i \le n-\ell} \min_{i \le j \le i+\ell} S_{j,n}$. Exact and weak limit results are obtained for the $M_{\ell,n}$.
Further we show that with probability one

$$\lim_{n\to\infty} \frac{(\ell+1)nM_{\ell,n}}{\log n} = 1 .$$

This extends results of Cheng.

Accession For	
NTIS GRA&I	✓
DTIC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

DTIC COPY INSPECTED

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

WEAK AND STRONG LAW RESULTS FOR A FUNCTION OF THE SPACINGS

William P. McCormick

Abstract

Let $\{U_n, n\geq 1\}$ be i.i.d. uniform on $(0,1)$ random variables and define $S_{i,n} = U_{i,n-1} - U_{i-1,n-1}$, $i=1,\ldots, n$ where the $U_{i,n-1}$ are the order statistics from a sample of size n-1 and $U_{0,n-1} = 0$ and $U_{n,n-1} = 1$. The $S_{i,n}$ are called the spacings divided by $U_1,\ldots, U_{n-1}$. For a fixed integer ℓ, set $M_{\ell,n} = \max_{1\leq i\leq n-\ell} \min_{i\leq j\leq i+\ell} S_{j,n}$. Exact and weak limit results are obtained for the $M_{\ell,n}$. Further we show that with probability one

$$\lim_{n\to\infty} \frac{(\ell+1)nM_{\ell,n}}{\log n} = 1 .$$

This extends results of Cheng.

Keywords: Order statistic, spacings, limiting distribution, strong law.

Supported by AFOSR contract F49620-82C-0009

Research was supported by the National Science Foundation under Grant MCS8202259.

1. Exact distribution.

Let $X_1, X_2, \ldots, X_n$ be i.i.d. having continuous distribution F. Let ℓ be a fixed intege~ and define a random variable $Y_{\ell,n} = \max_{1 \le i \le n-\ell} \min_{i \le j \le i+\ell} X_j$. In this paper we will determine the exact and limiting distribution of $Y_{\ell,n}$. In section 3 these results are then applied to obtain weak and strong laws for spacings generalizing previous work of Cheng [5]. Further these results may be of independent interest and we mention [4] in which a similar analysis has been carried out.

Let $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$ denote the order statistics and define the random index R_n by $Y_{\ell,n} = X_{R_{n,n}}$. If $r_1, r_2, \ldots, r_n$ denote the ranks of $X_1, X_2, \ldots, X_n$, it is clear that

$$R_n = \max_{1 \le i \le n-\ell} \min_{i \le j \le i+\ell} r_j \ . \tag{1.1}$$

Observe that R_n is independent of $X_{1,n}, \ldots, X_{n,n}$ and R_n has the distribution of the permutation statistic defined by the right hand side of (1.1) with all permutations equally likely. In the following we take R_n as defined on the space of permutations of $1,2,\ldots,n$. Then $Y_{\ell,n} \overset{d}{=} X_{R_{n,n}}$, that is, we have equality in distribution.

We introduce some convenient terminology. Define an r-component of a permutation as a collection of consecutive entries each of which is greater than r and the collection is maximal with respect to this property. The size of an r-component is defined to be the number of elements in the component. Further let $\beta_{j_1,\ldots,j_\ell} = \beta^{r,n}_{j_1,\ldots,j_\ell}$ equal the number of permutations on n elements with exactly j_k r-components of size k and no r-component of size greater than ℓ. Note

$$\#(R_n \le r) = \sum_{\underline{j} \in I_{n-r}} \beta_{j_1,\ldots,j_\ell} \tag{1.2}$$

where # denotes cardinality and $I_{n-r} = \{\underline{j} = (j_1,\ldots,j_\ell) : \sum_{t=1}^{\ell} t j_t = n - r$ and j_t is a nonnegative integer$\}$.

$\beta_{j_1,\ldots,j_\ell}$ may be evaluated by the following elementary counting argument. First select out of $\sum_1^\ell j_t$ places j_1 places for 1-components, j_2 places for 2-components,..., j_ℓ places for ℓ-components. This is done in $(\sum_1^\ell j_t)![\Pi_1^\ell j_t!]^{-1}$ ways. Next arrange the numbers r+1,..., n in one of (n-r)! ways and the numbers 1,..., r in one of r! ways. Finally choose $\sum_1^\ell j_t$ spaces among the r+1 spaces separating the numbers 1,..., r. This is done in $\binom{r+1}{\sum_1^\ell j_t}$ ways.

Notice that a permutation counted by $\beta_{j_1,\ldots,j_\ell}$ can be constructed as follows. Designate the spaces chosen in the last step as being a 1-component, 2-component,..., or ℓ-component according to the selection made in step 1. In these spaces make the appropriately sized component by placing the numbers r+1, ..., n according to the order given them by their permutation. Between these components place the numbers 1,..., r in the order given them by their permutation. This construction is also reversible. Hence

$$\beta_{j_1,\ldots,j_\ell} = \frac{(j_1+j_2+\ldots+j_\ell)!}{j_1!j_2!\ldots j_\ell!}(n-r)!r!\binom{r+1}{j_1+\ldots+j_\ell}. \tag{1.3}$$

Therefore by (1.2) and (1.3) we have proved the following.

Theorem 1.1

Let $X_1, X_2, \ldots, X_n$ be i.i.d. with continuous distribution F. Then

$$P\{Y_{\ell,n} \le x\} = \sum_{r=1}^n P(X_{r,n} \le x)P(R_n = r)$$

where

$$P\{R_n \le r\} = \frac{(n-r)!r!(r+1)!}{n!} \sum_{\underline{j}\in I_{n-r}} \frac{1}{\Pi_1^\ell j_t!(r+1-\Sigma_1^\ell j_t)!}, I_{n-r} = \{\underline{j}: \sum_1^\ell tj_t = n-r\}. \tag{1.4}$$

Remark: Note Theorem 1.1 remains true if the X_i are assumed only to be exchangeable.

The distribution of $Y_{\ell,n}$ can be obtained in another way which yields a simpler expression than that given in Theorem 1.1. For $A \subset \{1,2,\ldots, n\}$ let $M(A) = \max\{X_i, i\in A\}$ and $W(A) = \min\{X_i, i\in A\}$ with the convention $W(\emptyset) = \infty$. Let $E_{k,\ell}$

equal the class of all k element subsets of $\{1,2,\ldots, n\}$ which do not contain an interval of length greater than ℓ. Then

$$(1.5)\quad P\{Y_{\ell,n} \le x\} = \sum_{k=0}^{n} \sum_{A\in E_{k,\ell}} P\{W(A) > x,\ M(A^c) \le x\} = \sum_{k=0}^{n} \#(E_{k,\ell})(1 - F(X))^k F^{n-k}(x)\ .$$

To evaluate $\#(E_{k,\ell})$ we partition $E_{k,\ell}$ into the following sets. Let $\beta_{j_1,\ldots,j_\ell}$ be the class of all k element subsets of $\{1,\ldots, n\}$ containing j_i intervals of length i and no interval of length greater than ℓ. Then

$$(1.6)\qquad \#(E_{k,\ell}) = \sum_{\underline{j}\in I_k} \#(\beta_{j_1,\ldots,\ j_\ell}).$$

$\#(\beta_{j_1,\ldots,\ j_\ell})$ is obtained by the following counting argument. Consider n-k blocks into which integers will be put and the n-k+1 spaces between the blocks. Among these n-k+1 spaces choose j_1 to be designated as a single element space, j_2 for a two element space,..., j_ℓ for an ℓ-element space. Then a k element subset of $\{1,\ldots, n\}$ belonging to $\beta_{j_1,\ldots,\ j_\ell}$ is obtained by writing the numbers 1 to n in their natural order putting one integer in each of the n-k blocks and j_i consecutive integers in a space designated as an i-element space. The k-element set is then obtained by choosing the numbers put into the spaces. Hence

$$(1.7)\qquad \#(\beta_{j_1,\ldots,\ j_\ell}) = \frac{(n-k+1)!}{\Pi_1^\ell\ j_t!(n-k+1-\Sigma_1^\ell\ j_t)!}\ .$$

<u>Theorem 1.2</u>

Under the assumptions of Theorem 1.1 we have

$$(1.8)\qquad P\{Y_{\ell,n} \le x\} = \sum_{k=0}^{n} \sum_{\underline{j}\in I_k} \frac{(n-k+1)!(1-F(x))^k F^{n-k}(x)}{\Pi_1^\ell\ j_t!(n-k+1-\Sigma_1^\ell\ j_t)!}\ .$$

If the X_i are assumed only to be exchangeable and $F^{(k)}(x) = P\{W(A)>x, M(A^c)\le x\}$ where $A \subset \{1,\ldots, n\}$ is any k element subset then

$$P\{Y_{\ell,n} \le x\} = \sum_{k=0}^{n} \sum_{\underline{j}\in I_k} \frac{(n-k+1)!F^{(k)}(x)}{\Pi_1^{\ell} j_t!(n-k+1-\Sigma_1^{\ell} j_t)!} \quad . \tag{1.9}$$

Proof: (1.8) is immediate from (1.5), (1.6), and (1.7) while (1.9) follows for the same reasons except that in (1.5) the expression $(1-F(x))^k F^{n-k}(x)$ is replaced by $F^{(k)}(x)$.

2. Limiting distribution

In this section we derive the asymptotic behavior of $Y_{\ell,n}$. Preliminary to this work we obtain an asymptotic result for the permutation statistic R_n. In our analysis we rely on a method for obtaining the asymptotic behavior of sums with positive terms. A description of this tool may be found in the expository paper [3].

Lemma 2.1

Let R_n be the permutation statistic defined in (1.1) and having distribution given in (1.4). Then

$$\lim_{n\to\infty} P\left\{\frac{n-R_n}{n^{\ell/\ell+1}} \le x\right\} = \begin{cases} 0, & x<0 \\ 1-e^{-x^{\ell+1}}, & x\ge 0 \end{cases} \quad . \tag{2.1}$$

Proof: In order to obtain the asymptotic behavior of the sum in (1.4) we first locate the maximum summand and introduce a change of variables so that the largest term occurs at the zero point.

Observe that if $j_1^*,\ldots, j_\ell^*$ are defined as the solution to the equations

$$\begin{aligned} & j_1^* + 2j_2^* + \ldots + \ell j_\ell^* = n - r \\ & (j_1^*)^t = j_t^*(r - j_1^* - \ldots - j_\ell^*)^{t-1} \ , \ t=2,3,\ldots, \ell \end{aligned} \tag{2.2}$$

then the maximum summand in (1.4) occurs in a suitable neighborhood of $(j_1^*,\ldots,j_\ell^*)$. Let $h_i = (j_i^*)^{1/2}$, $i=2,\ldots, \ell$ and $j_i = j_i^* + x_i h_i$, $i=2,\ldots, \ell$ where x_i is a fractional index with step size $(h_i)^{-1}$. Making the change of variables in (1.4) and restricting attention to a neighborhood of the maximum summand, we consider

$$(2.3)\qquad \frac{(n-r)!r!(r+1)!}{n!}\sum{}^{*}\Big[\prod_{2}^{\ell}(j_t^*+x_th_t)!\,(n-r-\sum_{2}^{\ell}tj_t^*-\sum_{2}^{\ell}tx_th_t)!\Big]^{-1}$$

$$\cdot\Big[(r+1-\sum_{2}^{\ell}(t-1)j_t^*+\sum_{2}^{\ell}(t-1)x_th_t)!\Big]^{-1}$$

where the summation is over the x_i such that x_i has step size h_i^{-1} and $\max_{2\le i\le\ell}|x_i|\le A$ where A is a fixed positive constant. Then with $\max_{2\le i\le\ell}|x_i|\le A$, Stirling's formula and (2.2) we have

$$j_t! = (j_1^*)^{tj_t}(r-j_1^*-\dots-j_\ell^*)^{-(t-1)j_t}\,h_t\sqrt{2\pi}$$

$$(1+0(\frac{1}{j_t^*}))\exp\{j_t[\ell n(1+\frac{x_t}{h_t})-1]\}\quad,\quad t=2,\dots,\ell.$$

Similarly

$$j_1! = (j_1^*)^{j_1}\sqrt{2\pi j_1^*}(1+0(\frac{1}{j_1^*}))\exp\{j_1[\ell n(1-\frac{1}{j_1^*}\sum_{2}^{\ell}tx_th_t)-1]\}\ .$$

Therefore using (2.2) and $\ell n(1+x) = x-\frac{x^2}{2}+0(x^3)$ as $x\to 0$ it can be checked that

$$\prod_{1}^{\ell} j_t! = (j_1^*)^m\exp\{\tfrac{1}{2}\sum_{2}^{\ell}x_t^2-\sum_{1}^{\ell}j_t^*+0(\frac{1}{h_\ell}+\frac{j_2^*}{j_1^*})\}$$

$$\prod_{t=2}^{\ell}\Big[(r-\sum_{t=1}^{\ell}j_t^*)^{-(t-1)j_t}\,h_t\sqrt{2\pi}\Big]\sqrt{2\pi j_1^*}(1+0(\frac{1}{j_\ell^*}))$$

where $m = n-r$.

Therefore we find the expression at (2.3) equals

$$(2.4)\qquad \exp\{m\ell n(\frac{m}{j_1^*})+2r\ell n(1+\frac{\sum_1^\ell j_t^*}{r-\sum_1^\ell j_t^*})+n\ell n(1-\frac{m+\sum_1^\ell j_t^*}{n})\}(1+0(\frac{1}{h_\ell}+\frac{j_2^*}{j_1^*}))$$

$$\frac{1}{\prod_2^\ell\sqrt{2\pi}\,h_t}\sum{}^{*}e^{-\frac{1}{2}\sum_2^\ell x_t^2}\ .$$

Since the x_i have step size h_i^{-1} we have

$$\frac{1}{\prod_2^\ell\sqrt{2\pi}h_t}\sum{}^{*}e^{-\frac{1}{2}\sum_2^\ell x_t^2}\to\Big[\int_{-A}^{A}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\Big]^{\ell-1}\qquad\text{as } n\to\infty$$

so that it suffices to consider

$$\exp\{m\ell n(\frac{m}{j_1^*}) + 2r\ell n(1 + \frac{\Sigma_1^{\ell} j_t^*}{r-\Sigma_1^{\ell} j_t^*}) + n\ell n(1 - \frac{m+\Sigma_1^{\ell} j_t^*}{n})\} \ . \tag{2.5}$$

Let $x = \dfrac{j_1^*}{r-\Sigma_1^{\ell} j_t^*}$ and $\varepsilon = \dfrac{m}{r-\Sigma_1^{\ell} j_t^*}$. Then by (2.2) we have

$$\sum_1^{\ell} tx^t = \varepsilon \ . \tag{2.6}$$

For $f(x) = \sum_{k=0}^{K} a_k x^k + R(x)$ with $R(x) = O(x^{K+1})$ as $x \to 0$, let $\lceil f(x) \rceil_r = \sum_{k=0}^{r} a_k x^k$, $0 \le r \le K$. Using (2.6) we observe

$$\frac{m}{j_1^*} = \frac{\varepsilon}{x} = \sum_{t=0}^{\ell+1} (t+1)x^t - (\ell+1)x^{\ell} - (\ell+2)x^{\ell+1} = \lceil (1-x)^{-2} \rceil_{\ell+1} - (\ell+1)x^{\ell} - (\ell+2)x^{\ell+1} . \tag{2.7}$$

$$(1 + \frac{\Sigma_1^{\ell} j_t^*}{r-\Sigma_1^{\ell} j_t^*})^2 = (1 + \Sigma_1^{\ell} x^t)^2 = \lceil (1-x)^{-2} \rceil_{\ell+1} - 2x^{\ell+1} + O(x^{\ell+2}) \ . \tag{2.8}$$

$$1 - \frac{m+\Sigma_1^{\ell} j_t^*}{n} = \lceil 1+\varepsilon+\Sigma_1^{\ell} x^t \rceil^{-1} = (\sum_{t=0}^{\ell} (t+1)x^t)^{-1} = (\lceil (1-x)^{-2} \rceil_{\ell+1} - (\ell+2)x^{\ell+1})^{-1} \ . \tag{2.9}$$

Therefore by (2.7), (2.8), and (2.9) we see that (2.5) can be written

$$\begin{aligned}
&\exp\{m\ell n(\lceil (1-x)^{-2} \rceil_{\ell+1} - (\ell+1)x^{\ell} - (\ell+2)x^{\ell+1}) \\
&+ r\ell n(\lceil (1-x)^{-2} \rceil_{\ell+1} - 2x^{\ell+1} + O(x^{\ell+2})) \\
&- n\ell n(\lceil 1-x)^{-2} \rceil_{\ell+1} - (\ell+2)x^{\ell+1})\} \\
&= \exp\{-(\ell+1)m(\frac{j_t^*}{r-\Sigma_1^{\ell} j_t^*})^{\ell} - 2r(\frac{j_1^*}{r-\Sigma_1^{\ell} j_t^*})^{\ell+1} \\
&+ n(\ell+2)(\frac{j_1^*}{r-\Sigma_1^{\ell} j_t^*})^{(\ell+1)} + mO((\frac{j_1^*}{r-\Sigma_1^{\ell} j_t^*})^{(\ell+1)}) \\
&+ ([illegible]((\frac{j^*}{r-\Sigma_1 j_t^*})^{(\ell+2)})\} \ .
\end{aligned} \tag{2.10}$$

By taking $m = xn^{\frac{\ell}{\ell+1}}$ we have that $m \sim j_1^*$, $r \sim n$, as $n \to \infty$. Hence (2.10) is asymptotic to

$$\exp\{-x^{\ell+1} + 0(n^{-\frac{1}{\ell+1}})\} \tag{2.11}$$

Finally outside the neighborhood of the maximum summand we have that

$$\frac{(n-r)!r!(r+1)!}{n!} \sum^{**} [\prod_{t=2}^{\ell} (j_t^*+x_t h_t)!(n-r-\sum_2^{\ell} tj_t^* - \sum_2^{\ell} tx_t h_t)!]^{-1}$$

$$\cdot [(r+1-\sum_2^{\ell}(t-1)j_t^* + \sum_2^{\ell}(t-1)x_t h_t)!]^{-1}$$

$$= e^{-x^{\ell+1}}[1-[\int_{-A}^{A} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx]^{(\ell-1)}]0(1) \text{ as } n \to \infty \tag{2.12}$$

where $\sum^{**}$ denotes the summation over the x_i such that x_i has step h_i^{-1} and $\max_{2\le i\le \ell} |x_i| > A$. Hence Lemma 2.1 holds by (2.11) and (2.12).

Lemma 2.1 and results of Balkema and de Haan [1] and [2] allow us information on the asymptotic behavior of $Y_{\ell,n}$ even when classical extreme value theory does not apply. However, we must allow random normalization in the following weak limit results. Since the proofs of Theorem 2.1 and Theorem 2.2 are the same as the proofs of Proposition 6 and Proposition 8 of [4], we give only statements of the Theorems.

Theorem 2.1

Let $X_1, X_2, \ldots$ be an i.i.d. sequence with distribution F satisfying

(i) $F(x) < 1 \quad , \quad -\infty<x<\infty$

(ii) $\dfrac{F(ru)-F(u)}{(1-F(u))^{\frac{\ell+2}{2}}} \to \lambda \log r \quad , \quad 0<r<\infty$

as $u \to \infty$ for some $\lambda > 0$. Let $Z_n = \dfrac{X_{R_{n,n}}}{f(\frac{R_n}{n})}$ where f denotes the right continuous inverse of F. Then

$$\lim_{n\to\infty} P\{Z_n \le z\} = \int_0^\infty \Phi(\lambda w^{\frac{\ell+1}{2}} \ell nz)(\ell+1)w^{\ell} e^{-w^{\ell+1}}\, dw \quad , \quad 0<z<\infty$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du$.

Remark: It can be checked that the hypothesis of Theorem 2.1 is satisfied for F of the form $F(x) = 1 - c(\ell nx)^{-(2/\ell)}$ for x large. By classical results in extreme value theory, such distributions do not belong to the domain of attraction of any extreme value distribution.

Theorem 2.2

Let F be a distribution function with upper endpoint y_1 which has a strictly positive density $F'(y) = e^{u(y)}$ for y in a left neighborhood of y_1. Suppose F satisfies one of the following conditions:

(i) $1 - F(y') \sim 1 - F(y)$ for $y \to y_1$ implies $F'(y') \sim F'(y)$,

(ii) $\frac{F''(y)(1-F(y))}{(F'(y))^2}$ is bounded in a left neighborhood of y_1,

(iii) F' varies regularly in y_1- with exponent $p \neq -1$,

(iv) $\limsup_{y\to y_1-} \frac{u''(y)}{(u'(y))^2} < 1$ or $\liminf_{y\to y_1-} \frac{u''(y)}{(u'(y))^2} > 1$.

Then if f denotes the right continuous inverse of F, we have

$$\lim_{n\to\infty} P\left\{\frac{Y_{\ell,n} - f(\frac{R_n}{n})}{f(\frac{R_n}{n} + \frac{R_n(n-R_n)}{n^3}) - f(\frac{R_n}{n})} \le x\right\} = \Phi(x) \quad , \quad -\infty<x<\infty .$$

Remark: The conditions of Theorem 2.2 are satisfied for the normal, Laplace, Cauchy, beta, gamma distributions and all limit distributions for extreme order statistics $X_{k,n}$ with k fixed or n-k fixed.

Let $V_i = \min_{i\le j\le i+\ell} X_j$, $i=1,\dots, n-\ell$. Then $Y_{\ell,n} = \max_{1\le i\le n-\ell} V_i$ and the V_i form an ℓ-dependent stationary sequence, that is V_i and V_j are independent if $|i-j| > \ell$.

Extreme value theory for dependent stationary sequences has received considerable attention. A basic result which relates the asymptotic behavior of the maximum of a stationary sequence to the asymptotic behavior of the maximum of the associated i.i.d. sequence is the following:

Let $\{\xi_n, n\geq 1\}$ be a stationary sequence and let $F_{i_1,\ldots,i_n}(x) = P\{\xi_{i_1}\leq x,\ldots,\xi_{i_n}\leq x\}$. Condition $D(u_n)$ is said to hold if for any integers $1\leq i_1<\ldots<i_p<j_1<\ldots<j_q\leq n$ with $j_1 - i_p \geq \ell$ we have

$$|F_{i_1\ldots i_p j_1\ldots j_q}(u_n) - F_{i_1\ldots i_p}(u_n)F_{j_1\ldots j_q}(u_n)| \leq \alpha_{n,\ell}$$

where $\alpha_{n,\ell_n} = o(1)$ for some sequence $\ell_n = o(n)$. Condition $D'(u_n)$ is said to hold for the stationary sequence $\{\xi_n, n\geq 1\}$ if

$$\overline{\lim_{n\to\infty}}\, n \sum_{j=2}^{[n/k]} P\{\xi_1>u_n, \xi_j>u_n\} \to 0 \text{ as } k\to\infty.$$

In [7] it is shown (Theorem 3.5.2) that if $M_n = \max_{1\leq i\leq n} \xi_i$ and $\hat{M}_n = \max_{1\leq i\leq n} \hat{\xi}_i$ where the $\hat{\xi}_i$ are i.i.d. with the same underlying distribution as the ξ_i then if for every x conditions $D(u_n)$ and $D'(u_n)$ hold where $u_n = x/a_n + b_n$, $a_n > 0$, we have

$$\lim_{n\to\infty} P\{M_n \leq \frac{x}{a_n} + b_n\} = G(x) \quad , \quad -\infty<x<\infty$$

if and only if

$$\lim_{n\to\infty} P\{\hat{M}_n \leq \frac{x}{a_n} + b_n\} = G(x) \quad , \quad -\infty<x<\infty .$$

Therefore in view of the above result the asymptotic behavior of the max v_i will be exactly the same as if the v_i were independent once $D(u_n)$ and $D'(u_n)$ have been established where u_n is determined by $(1-F(u_n))^{\ell+1} = \frac{\tau}{n} + 0(\frac{1}{n})$ as $n\to\infty$ where F is the underlying distribution of the X_i. Such a sequence $\{u_n, n\geq 1\}$ exists for F belonging to the domain of attraction of an extreme value distribution. Finally the verification of Conditions $D(u_n)$ and $D'(u_n)$ is easy.

In order to establish a strong law result in section 3 we need the following result.

Theorem 2.3

Let $\{X_n, n\geq 1\}$ be an i.i.d. sequence with underlying distribution F. Let $Y_{\ell,n} = \max_{1\leq i\leq n-\ell} \min_{i\leq j\leq i+\ell} X_j$. If x_n is any sequence such that $n(1-F(x_n))^{\ell+2} = 0(1)$ then

$$P\{Y_{\ell,n} \leq x_n\} = (1+0([n(1-F(x_n))^{\ell}]^{-1})$$
$$+ 0(n[1-F(x_n)]^{\ell+2})\exp\{-n(1-F(x_n))^{\ell+1}\} .$$

Proof: Since the technique to obtain the asymptotic behavior of $Y_{\ell,n}$ is the same as the one used for R_n, we present a sketch only. In order to obtain the asymptotic behavior for

$$H(x) = \sum_{(\underline{j},r)\in J} \frac{(r+1)!(1-F(x))^{n-r}(F(x))^r}{\Pi j_t!(r+1-\Sigma_1^{\ell} j_t)!}$$

where $J = \{(\underline{j},r): 0\leq j_t, r, t=1,2,\ldots,\ell \text{ and } \Sigma_1^{\ell} t j_t = n-r\}$, we find that the maximum summand occurs in a neighborhood of the point $(\underline{j}^*,r^*)$ satisfying

$$\text{(i)} \quad \sum_1^{\ell} t j_t^* = n - r^*$$

$$(2.13) \qquad \text{(ii)} \quad (j_1^*)^t = (j_t^*)(r^* - \sum_1^{\ell} j_t^*)^{(t-1)}, \quad t=2,\ldots,\ell$$

$$\text{(iii)} \quad \alpha r^* j_1^* = (r^* - \sum_1^{\ell} j_t^*)^2$$

where $\alpha = F(x)(1-F(x))^{-1}$.

By the usual calculations we find that

$$(2.14) \qquad H(x) = (\frac{\alpha}{1+\alpha})^n (\frac{r}{\alpha j_1^*})^{n/2} (1+0(\frac{\alpha^{\ell}}{n})) .$$

Let $u = \alpha j_1^*/r^*$. Then by (2.13 ii) we find that $j_t^*/j_1^* = (u/\alpha)^{(t-1)}$ so that from (2.13 iii) $u^{-2} = (u^{-2} - \alpha^{-1}\Sigma_{t=0}^{\ell-1} (u/\alpha)^t)^2$ which implies $(1+\alpha^{-1})u - (u/\alpha)^{\ell+1} + 0(\alpha^{-(\ell+2)}) = 1$. Hence

$$(2.15) \qquad u = \frac{\alpha}{\alpha+1} + (1+\alpha)^{-(\ell+1)} + 0(\alpha^{-(\ell+2)}) \ .$$

By (2.14) and (2.15) we find

$$H(x) = (1 + 0(n^{-1}\alpha^{\ell}))$$

$$\exp\{n\ell n(\frac{\alpha}{1+\alpha}) - n\ell n(\frac{\alpha}{1+\alpha} + (1+\alpha)^{-(\ell+1)} + 0(\alpha^{-(\ell+2)}))\}$$

$$= (1 + 0(n^{-1}\alpha^{\ell}) + 0(n\alpha^{-(\ell+2)}))\exp\{-n(1+\alpha)^{-(\ell+1)}\}$$

proving Theorem 2.3.

3. Weak and Strong Laws for Spacings.

In this sectic we derive weak and strong law results for a particular function of the spacings. Let $\{U_n, n\geq 1\}$ be i.i.d. uniform on $(0,1)$ random variables. Let $U_{1,n}\leq\ldots\leq U_{n,n}$ be the order statistics for $U_1,\ldots, U_n$. Then the random variables $S_{i,n+1} = U_{i,n} - U_{i-1,n}$, $i=1,\ldots, n+1$ are called the spacings divided by $U_1,\ldots,U_n$ where $U_{0,n} = 0$ and $U_{n+1,n} = 1$.

Let $M_{\ell,n+1} = \max\limits_{1\leq i\leq n+1-\ell} \min\limits_{i\leq j\leq i+\ell} S_{j,n+1}$. The quantity $M_{1,n}$ played a role in the work of Marron [8] and Chow, Geman and Wu [6] in cross-validated kernel density estimation. Statistical properties of $M_{1,n}$ were analyzed in Cheng [5].

Our first result gives the exact distribution of $M_{\ell,n}$.

Theorem 3.1

$$P\{M_{\ell,n} \leq x\} = \sum_{k=0}^{n} \sum_{\underline{j}\in I_k} \frac{(n-k+1)!}{\Pi_1^{\ell} j_t!(n+1-k-\Sigma_1^{\ell} j_t)!} \sum_{t=0}^{n-k} (-1)^t \binom{n-k}{t} \{[1-(k+t)x]_+\}^{n-1}$$

where $x_+ = x$ if $x > 0$ or $= 0$ if $x \leq 0$.

Proof: The result follows from Theorem 1.2 and the fact that

$$P\{S_{1,n} > x,\ldots, S_{k,n} > x, S_{k+1,n} \leq x,\ldots, S_{n,n} \leq x\}$$

$$= \sum_{t=0}^{n-k} (-1)^t \binom{n-k}{t} \{[1-(k+t)x]_+\}^{n-1}$$

which was shown in [5], p. 3.

To obtain an asymptotic result for $M_{\ell,n}$ the following representation will be useful.

<u>Lemma 3.1</u> (Pyke) Let $\{X_n, n\geq 1\}$ be i.i.d. exponential random variable with mean one. Let $T_n = \Sigma_{i=1}^n X_i$. Then

$$(S_{1,n}, S_{2,n},\ldots, S_{n,n}) \overset{d}{=} (X_1/T_n, X_2/T_n,\ldots, X_n/T_n).$$

<u>Theorem 3.2</u>

$$\lim_{n\to\infty} P\{M_{\ell,n} \leq \frac{x+\ell nn}{n(\ell+1)}\} = \exp(-e^{-x}) , \quad -\infty<x<\infty .$$

<u>Proof</u>: Let $Z_{\ell,n} = \max_{1\leq i\leq n-\ell} \min_{i\leq j\leq i+\ell} X_j$ where the X_i are i.i.d., $X_i \sim 1 - e^{-x}$, and let $T_n = \Sigma_1^n X_i$. Then by Lemma 3.1 we have

$$P\{M_{\ell,n} \leq \frac{x+\ell nn}{(\ell+1)n}\} = P\{\frac{Z_{\ell,n}}{T_n} \leq \frac{x+\ell nn}{(\ell+1)n}\} .$$

But $((T_n/n) - 1)\ell nn \xrightarrow{p} 1$ as $n \to \infty$. See [5], p. 7. Therefore it suffices to consider

$$P\{Z_{\ell,n} \leq (x + \ell nn)\frac{1}{\ell+1}\} .$$

Since the variables $W_i = \min_{i\leq j\leq i+\ell} X_i$ satisfy $D(u_n)$ and $D'(u_n)$ with u_n given by $P\{W_i > u_n\} = e^{-(\ell+1)u_n} = \frac{e^{-x}}{n}$, we have that the limiting distribution of $Z_{\ell,n}$ is the same as if the variables W_i were independent. Hence the Theorem holds. Alternatively we may apply Theorem 2.3 with $x_n=(1/(\ell+1))\,(x+\ell nn)$.

Finally we conclude with a strong law result for the spacings. With Theorem 2.3 established it is easy to check that the method of proof of Theorem 4.8 in [5] carries over with obvious modifications to the present case. Therefore, we only state the result

<u>Theorem 3.3</u>

With probability one

$$\lim_{n\to\infty} \frac{(\ell+1)nM_{\ell,n}}{\log n} = 1 .$$

REFERENCES

[1] Balkema, A.A. and de Haan, L. (1978). Limit distributions for order statistics. I Theory Prob. Appl. 23, 77-92.

[2] Balkema, A.A. and de Haan, L. (1978). Limit distributions for order statistics. II Theory Prob. Appl. 23, 341-358.

[3] Bender, E.A. (1974). Asymptotic methods in enumeration. Siam Review 16, 485-515.

[4] Canfield, E.R. and McCormick, W.P. (1982). Exact and limiting distributions for sustained maxima

[5] Cheng, S. (1983). On a problem concerning spacings. Center for Stochastic Processes Tech. Rept. #27, Univ. of North Carolina, Chapel Hill, N.C.

[6] Chow, Y.S., Geman, S., and Wu, L.D. (1983). Consistent cross-validated density estimation. Ann. Statist. 11, to appear.

[7] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). Extremes and Related Properties of Random Sequences and Processes, Springer, N.Y.

[8] Marron, J.S. (1983). As asymptotic efficient solution to the bandwidth problem of kernel density estimation. Inst. Stat. Mimeo Series #1518, Univ. of North Carolina, Chapel Hill, N.C.

[9] Pyke, R. (1965). Spacings, J. Roy. Statist. Soc. 27, 395-436.

END

DATE
FILMED

8 — 83

DTIC